

RECEIVED

AUG 01 2002

TECH CENTER 1600/2900



8/7/02 1600

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/586,106B

DATE: 07/24/2002  
 TIME: 11:51:58

Input Set : A:\08411-032001.txt  
 Output Set: N:\CRF3\07242002\I586106B.raw

p.6

3 <110> APPLICANT: Wright, David A.  
 4 Voytas, Daniel F.  
 6 <120> TITLE OF INVENTION: Plant Retroelements and Methods Related Thereto  
 8 <130> FILE REFERENCE: P-1065A  
 10 <140> CURRENT APPLICATION NUMBER: 09/586,106B  
 11 <141> CURRENT FILING DATE: 2000-06-02  
 13 <150> PRIOR APPLICATION NUMBER: 60/087,125  
 14 <151> PRIOR FILING DATE: 1998-05-29  
 16 <150> PRIOR APPLICATION NUMBER: 09/322,478  
 17 <151> PRIOR FILING DATE: 1999-05-28  
 19 <160> NUMBER OF SEQ ID NOS: 169  
 21 <170> SOFTWARE: PatentIn Ver. 2.1  
 23 <210> SEQ ID NO: 1  
 24 <211> LENGTH: 18  
 25 <212> TYPE: DNA  
 26 <213> ORGANISM: Glycine max  
 28 <400> SEQUENCE: 1  
 29 tggcgccggtt gccaatg 18  
 32 <210> SEQ ID NO: 2  
 33 <211> LENGTH: 18  
 34 <212> TYPE: DNA  
 35 <213> ORGANISM: Glycine max  
 37 <400> SEQUENCE: 2 18  
 38 tggcgccggtt gtcgggga  
 41 <210> SEQ ID NO: 3  
 42 <211> LENGTH: 6  
 43 <212> TYPE: DNA  
 44 <213> ORGANISM: Glycine max  
 46 <400> SEQUENCE: 3 6  
 47 ttggggg  
 50 <210> SEQ ID NO: 4  
 51 <211> LENGTH: 7  
 52 <212> TYPE: PRT  
 53 <213> ORGANISM: Artificial Sequence  
 55 <220> FEATURE:  
 56 <223> OTHER INFORMATION: Description of Artificial Sequence: plant  
 57 retroelement sequence  
 59 <400> SEQUENCE: 4  
 60 Met Ala Ser Arg Lys Arg Lys  
 61 1 5  
 64 <210> SEQ ID NO: 5  
 65 <211> LENGTH: 1263  
 66 <212> TYPE: DNA

ENTERED

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/586,106B

DATE: 07/24/2002  
TIME: 11:51:58

Input Set : A:\08411-032001.txt  
Output Set: N:\CRF3\07242002\I586106B.raw

RECEIVED  
AUG 01 2002  
TECH CENTER 1600/2900

67 <213> ORGANISM: Artificial Sequence  
69 <220> FEATURE:  
70 <223> OTHER INFORMATION: Description of Artificial Sequence: plant  
71 retroelement sequence  
73 <400> SEQUENCE: 5  
74 atggcctccc gtaaacgcaa agctgtgccc acaccgggg aagcgtccaa ctgggactct 60  
75 tcacgtttca ctttcgagat tgcttggcac agataccagg atagcattca gctccggaac 120  
76 atccttccag agaggaatgt agagcttggg ccagggatgt ttgatgagtt cctgcaggaa 180  
77 ctccagaggc tcagatggga ccaggttctg acccgacttc cagagaagtg gattgatgtt 240  
78 gctctgggtga aggagtttta ctccaacctata tatgatccag aggaccacag tccgaagttt 300  
79 tggagtgttc gaggacaggt tgtgagattt gatgctgaga cgattaatga tttcctcgac 360  
80 accccgggtca tcttggcaga gggagaggat tatccagcct actctcagta cctcagcact 420  
81 cctccagacc atgatgccat cctttccgct ctgtgtactc cagggggacg atttgttctg 480  
82 aatgttgata gtgccccctg gaagctgctg cggaaggatc tgatgacgct cgcgcacaga 540  
83 tggagtgtgc tctcttattt taaccttgca ctgacttttc acacttctga tattaatgtt 600  
84 gacagggccc gactcaatta tggcttgggt atgaagatgg acctggacgt gggcagcctc 660  
85 atttctcttc agatcagtcg gatcgcccag tccatcactt ccaggcttgg gttcccagcg 720  
86 ttgatcacia cactgtgtga gattcagggg gttgtctctg ataccctgat ttttgagtca 780  
87 ctcatgctg tgatcaacct tgccctacatt aagaagaact gctggaacct tgccgatcca 840  
88 tctatcacat ttcaggggac ccgcccgcag cgcaccagag cttcggcgct ggcatctgag 900  
89 gctcctcttc catcccagca tcttctcag cctttttccc agagaccacg gcctccactt 960  
90 ctatccacct cagcacctcc atacatgcat ggacagatgc tcaggctcct gtaccagggt 1020  
91 cagcagatca tcattcagaa cctgtatcga ttgtccctac atttgagat ggatctgcca 1080  
92 ctcatgactc cggaggccta tcgtcagcag gtcgccaagc taggagacca gccctccact 1140  
93 gacagggggg aagagccttc tggagccgct gctactgagg atcctgccgt tgatgaagac 1200  
94 ctcatagctg acttggtctg cgctgattgg agcccatggg cagacttggg cagaggcagc 1260  
95 tga 1263

98 <210> SEQ ID NO: 6  
99 <211> LENGTH: 421  
100 <212> TYPE: PRT  
101 <213> ORGANISM: Artificial Sequence  
103 <220> FEATURE:  
104 <223> OTHER INFORMATION: Description of Artificial Sequence: plant  
105 retroelement sequence  
107 <400> SEQUENCE: 6  
108 Met Ala Ser Arg Lys Arg Lys Ala Val Pro Thr Pro Gly Glu Ala Ser  
109 1 5 10 15  
111 Asn Trp Asp Ser Ser Arg Phe Thr Phe Glu Ile Ala Trp His Arg Tyr  
112 20 25 30  
114 Gln Asp Ser Ile Gln Leu Arg Asn Ile Leu Pro Glu Arg Asn Val Glu  
115 35 40 45  
117 Leu Gly Pro Gly Met Phe Asp Glu Phe Leu Gln Glu Leu Gln Arg Leu  
118 50 55 60  
120 Arg Trp Asp Gln Val Leu Thr Arg Leu Pro Glu Lys Trp Ile Asp Val  
121 65 70 75 80  
123 Ala Leu Val Lys Glu Phe Tyr Ser Asn Leu Tyr Asp Pro Glu Asp His  
124 85 90 95  
126 Ser Pro Lys Phe Trp Ser Val Arg Gly Gln Val Val Arg Phe Asp Ala  
127 100 105 110

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/586,106B

DATE: 07/24/2002

TIME: 11:51:58

Input Set : A:\08411-032001.txt

Output Set: N:\CRF3\07242002\I586106B.raw

RECEIVED  
AUG 01 2002  
TECH CENTER 1600/2900

```

129 Glu Thr Ile Asn Asp Phe Leu Asp Thr Pro Val Ile Leu Ala Glu Gly
130      115      120      125
132 Glu Asp Tyr Pro Ala Tyr Ser Gln Tyr Leu Ser Thr Pro Pro Asp His
133      130      135      140
135 Asp Ala Ile Leu Ser Ala Leu Cys Thr Pro Gly Gly Arg Phe Val Leu
136 145      150      155      160
138 Asn Val Asp Ser Ala Pro Trp Lys Leu Leu Arg Lys Asp Leu Met Thr
139      165      170      175
141 Leu Ala Gln Thr Trp Ser Val Leu Ser Tyr Phe Asn Leu Ala Leu Thr
142      180      185      190
144 Phe His Thr Ser Asp Ile Asn Val Asp Arg Ala Arg Leu Asn Tyr Gly
145      195      200      205
147 Leu Val Met Lys Met Asp Leu Asp Val Gly Ser Leu Ile Ser Leu Gln
148      210      215      220
150 Ile Ser Gln Ile Ala Gln Ser Ile Thr Ser Arg Leu Gly Phe Pro Ala
151 225      230      235      240
153 Leu Ile Thr Thr Leu Cys Glu Ile Gln Gly Val Val Ser Asp Thr Leu
154      245      250      255
156 Ile Phe Glu Ser Leu Ser Pro Val Ile Asn Leu Ala Tyr Ile Lys Lys
157      260      265      270
159 Asn Cys Trp Asn Pro Ala Asp Pro Ser Ile Thr Phe Gln Gly Thr Arg
160      275      280      285
162 Arg Thr Arg Thr Arg Ala Ser Ala Ser Ala Ser Glu Ala Pro Leu Pro
163      290      295      300
165 Ser Gln His Pro Ser Gln Pro Phe Ser Gln Arg Pro Arg Pro Pro Leu
166 305      310      315      320
168 Leu Ser Thr Ser Ala Pro Pro Tyr Met His Gly Gln Met Leu Arg Ser
169      325      330      335
171 Leu Tyr Gln Gly Gln Gln Ile Ile Ile Gln Asn Leu Tyr Arg Leu Ser
172      340      345      350
174 Leu His Leu Gln Met Asp Leu Pro Leu Met Thr Pro Glu Ala Tyr Arg
175      355      360      365
177 Gln Gln Val Ala Lys Leu Gly Asp Gln Pro Ser Thr Asp Arg Gly Glu
178      370      375      380
180 Glu Pro Ser Gly Ala Ala Ala Thr Glu Asp Pro Ala Val Asp Glu Asp
181 385      390      395      400
183 Leu Ile Ala Asp Leu Ala Gly Ala Asp Trp Ser Pro Trp Ala Asp Leu
184      405      410      415
186 Gly Arg Gly Ser Glx
187      420
190 <210> SEQ ID NO: 7
191 <211> LENGTH: 1596
192 <212> TYPE: DNA
193 <213> ORGANISM: Artificial Sequence
195 <220> FEATURE:
196 <223> OTHER INFORMATION: Description of Artificial Sequence: plant
197      retroelement sequence
199 <400> SEQUENCE: 7
200 atgcgaggtga gaactgcac tggagacgtt gttcctatta acttagaaat tgaagctacg 60

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/586,106B

DATE: 07/24/2002

TIME: 11:51:58

Input Set : A:\08411-032001.txt

Output Set: N:\CRF3\07242002\I586106B.raw

RECEIVED  
AUG 01 2002  
TECH CENTER 1600/2900

```

201 tgtcggcgta acaacgctgc aagaagaaga agggagcaag acatagaagg aagtagttac 120
202 acctcacctc ctcttctctc aaattatgct cagatggacg gggaaccggc acaaagagtc 180
203 acactagagg acttctctaa taccaccact cctcagttct ttacaagtat cacaaggccg 240
204 gaagtccaag cagatctcct tactcaaggg aacctcttcc atggtcttcc aaatgaagat 300
205 ccatatgcgc atctagcctc atacatagag atatgcagca ccgttaaaat cgccggagtt 360
206 ccaaaagatg cgatactcct taacctcttt tccttttccc tagcaggaga ggcaaaaaga 420
207 tggttgcact cctttaaagg caatagctta agaacatggg aagaagtagt ggaaaaattc 480
208 ttaaagaagt atttcccaga gtcaaagacc gtcgaacgaa agatggagat ttcttatttc 540
209 catcaatttc tggatgaatc ccttagcgaa gcaactagacc atttccacgg attgctaaga 600
210 aaaacaccaa cacacagata cagcgagcca gtacaactaa acatattcat cgatgacttg 660
211 caactcttaa tcgaaacagc tactagaggg aagatcaagc tgaagactcc cgaagaagcg 720
212 atggagctcg tcgagaacat ggcgggctagc gatcaagcaa tccttcata tcacacttat 780
213 gttcccacaa aaagaagcct cttggagctt agcacgcagg acgcaacttt ggtacaaaac 840
214 aagctgttga cgaggcagat agaagccctc atcgaaaccc tcagcaagct gcctcaacaa 900
215 ttacaagcga taagttcttc ccactcttct gttttgcagg tagaagaatg ccccatatgc 960
216 agagggacac atgagcctgg acaatgtgca agccaacaag acccctctcg tgaagtaaatt 1020
217 tatataggca tactaaatcg ttacggattt cagggctaca accagggaaa tccatctgga 1080
218 ttcaatcaag gggcaacaag atttaatac gagccaccgg ggtttaatca aggaagaaac 1140
219 ttcattgcaag gctcaagttg gacgaataaa ggaaatcaat ataaggagca aaggaaccaa 1200
220 ccaccatacc agccaccata ccagcaccct agccaaggtc cgaatcagca agaaaagccc 1260
221 accaaaatag aggaactgct gctgcaattc atcaaggaga caagatcaca tcaaaagagc 1320
222 acggatgcag ccattcgga tctagaagtt caaatgggcc aactggcgca tgacaaagcc 1380
223 gaacggccca ctagaacttt cgggtgctaac atggagagaa gaaccccaag gaaggataaa 1440
224 gcagtactga ctagagggca gagaagagcg caggaggagg gtaagggtga aggagaagac 1500
225 tggccagaag aaggaaggac agagaagaca gaagaagaag agaagggtgc agaagaacct 1560
226 aagcgtacca agagccagag agcaagggaa gccaaag 1596

```

229 &lt;210&gt; SEQ ID NO: 8

230 &lt;211&gt; LENGTH: 532

231 &lt;212&gt; TYPE: PRT

232 &lt;213&gt; ORGANISM: Artificial Sequence

234 &lt;220&gt; FEATURE:

235 &lt;223&gt; OTHER INFORMATION: Description of Artificial Sequence: plant

236 retroelement sequence

238 &lt;400&gt; SEQUENCE: 8

239 Met Arg Gly Arg Thr Ala Ser Gly Asp Val Val Pro Ile Asn Leu Glu

240 1 5 10 15

242 Ile Glu Ala Thr Cys Arg Arg Asn Asn Ala Ala Arg Arg Arg Arg Glu

243 20 25 30

245 Gln Asp Ile Glu Gly Ser Ser Tyr Thr Ser Pro Pro Pro Ser Pro Asn

246 35 40 45

248 Tyr Ala Gln Met Asp Gly Glu Pro Ala Gln Arg Val Thr Leu Glu Asp

249 50 55 60

251 Phe Ser Asn Thr Thr Thr Pro Gln Phe Phe Thr Ser Ile Thr Arg Pro

252 65 70 75 80

254 Glu Val Gln Ala Asp Leu Leu Thr Gln Gly Asn Leu Phe His Gly Leu

255 85 90 95

257 Pro Asn Glu Asp Pro Tyr Ala His Leu Ala Ser Tyr Ile Glu Ile Cys

258 100 105 110

260 Ser Thr Val Lys Ile Ala Gly Val Pro Lys Asp Ala Ile Leu Leu Asn



RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/09/586,106B

DATE: 07/24/2002  
TIME: 11:51:59

Input Set : A:\08411-032001.txt  
Output Set: N:\CRF3\07242002\I586106B.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:166; N Pos. 6,15,16,18  
Seq#:168; N Pos. 7

RECEIVED  
AUG 01 2002  
TECH CENTER 1600/2900